
Article:
van Houdt, P.J. orcid.org/0000-0001-7431-8386, Ragunathan, S. orcid.org/0000-0002-2871-1758, Berks, M. orcid.org/0000-0003-4727-2006 et al. (12 more authors) (2024) Contrast-agent-based perfusion MRI code repository and testing framework: ISMRM Open Science Initiative for Perfusion Imaging (OSIPI). Magnetic Resonance in Medicine, 91 (5). pp. 1774-1786. ISSN 0740-3194

https://doi.org/10.1002/mrm.29826

Received: 24 March 2023 | Revised: 30 June 2023 | Accepted: 25 July 2023

DOI: 10.1002/mrm.29826

RESEARCH ARTICLE

Magnetic Resonance in Medicine

Contrast-agent-based perfusion MRI code repository and testing framework: ISMRM Open Science Initiative for Perfusion Imaging (OSIPI)

Petra J van Houdt[1] | Sudarshan Ragunathan[2] | Michael Berks[3] | Zaki Ahmed[4] | Lucy E Kershaw[5] | Oliver J Gurney-Champion[6,7] | Sirisha Tadimalla[8] | Jonathan Arvidsson[9,10] | Yu Sun[8] | Jesper Kallehauge[11,12] | Ben Dickie[13,14] | Simon Lévy[15] | Laura Bell[16] | Steven Sourbron[17] | Michael J Thrippleton[18] | on behalf of the ISMRM Perfusion Study Group

Correspondence
Michael J. Thrippleton, Centre for Clinical Brain Sciences, University of Edinburgh, The Chancellor's Building, 49 Little France Crescent, Edinburgh EH16 4SB, UK.
Email: m.j.thrippleton@ed.ac.uk

Abstract

Purpose: Software has a substantial impact on quantitative perfusion MRI values. The lack of generally accepted implementations, code sharing and transparent testing reduces reproducibility, hindering the use of perfusion MRI in clinical trials. To address these issues, the ISMRM Open Science Initiative for Perfusion Imaging (OSIPI) aimed to establish a community-led, centralized repository for sharing open-source code for processing contrast-based perfusion imaging, incorporating an open-source testing framework.

Methods: A repository was established on the OSIPI *GitHub* website. *Python* was chosen as the target software language. Calls for code contributions were made to OSIPI members, the ISMRM Perfusion Study Group, and publicly via OSIPI websites. An automated unit-testing framework was implemented to evaluate the output of code contributions, including visual representation of the results.

Results: The repository hosts 86 implementations of perfusion processing steps contributed by 12 individuals or teams. These cover all core aspects of DCE- and DSC-MRI processing, including multiple implementations of the same functionality. Tests were developed for 52 implementations, covering five analysis steps. For T_1 mapping, signal-to-concentration conversion and population AIF functions, different implementations resulted in near-identical output values. For the five pharmacokinetic models tested (Tofts, extended Tofts-Kety, Patlak, two-compartment exchange, and two-compartment uptake), differences in output parameters were observed between contributions.

For affiliations refer to page 10

Conclusions: The OSIPI DCE-DSC code repository represents a novel community-led model for code sharing and testing. The repository facilitates the re-use of existing code and the benchmarking of new code, promoting enhanced reproducibility in quantitative perfusion imaging.

KEYWORDS

dynamic contrast-enhanced MRI, dynamic susceptibility-contrast MRI, open source, OSIPI, perfusion, software

1 | INTRODUCTION

Perfusion imaging with DCE- or dynamic susceptibility-contrast (DSC-) MRI provides important pathophysiological information in several fields of clinical research, for example in neurodegenerative conditions, neurovascular disease, and oncology.[1–5] These techniques typically involve advanced or bespoke acquisition protocols, image analysis pipelines and software to generate quantitative or semi-quantitative perfusion parameters. A consequence is that perfusion and permeability parameters can vary widely between research groups and software implementations[3] and, while multisite reproducibility studies are unfortunately rare,[6] there is evidence to suggest that reproducibility is poor.[7] This impedes translation from single-site tools to harmonized quantitative imaging biomarkers for general use in clinical studies, multicenter trials, and for clinical diagnosis and monitoring.[6]

Several factors contribute to the variability, including differences in scanner hardware, pulse sequence, acquisition parameters and data processing pipelines. Initiatives such as the quantitative imaging biomarker alliance (QIBA),[8] and disease-specific initiatives including in neurodegenerative disease[4] and brain tumor imaging[9,10] have proposed recommendations for increased harmonization of acquisition and analysis in perfusion imaging. Image analysis software is increasingly recognized as a major determinant of the reliability and reproducibility of quantitative perfusion MRI parameters: several studies demonstrate substantial variation in results dependent on the software used to process DCE- or DSC-MRI data.[11–17] Researchers commonly develop or re-use in-house code due to the specialized nature of the processing or because existing available solutions are limited, unvalidated or difficult to use. However, most scientists are self-taught programmers[18,19] untrained in state-of-the art software development practices such as version control and unit testing, despite growing recognition that software "should be built, checked, and used as carefully as any physical apparatus".[20,21] Furthermore, sharing code in public repositories is not yet standard practice, but is important to improve the reproducibility of results.[22] Proprietary commercial software may reach a higher standard in software engineering terms, but the implementation details are hidden. Together, these factors lead to site-dependent results, increase the likelihood of errors, reduce transparency and impede replication.

Improvement in coding practices is essential but not sufficient to improve reliability and reproducibility of quantitative imaging biomarkers. Community-driven initiatives are also needed to ensure that code can be re-used *in practice*, to facilitate open testing using trusted reference data and comparison between implementations, and to develop community-led software libraries that become accepted standards or benchmarks within their fields. To address these challenges in the context of perfusion imaging, Taskforce 2.3 of the ISMRM Open Science Initiative for Perfusion Imaging (ISMRM OSIPI, referred to hereafter as "OSIPI")[23] was established with the following aims (Figure 1):

(i) To initiate a centralized repository for hosting open-source code that is maintained by the perfusion imaging community to promote code sharing, reduce the need for time-consuming duplicate development, facilitate reproducible analysis and make perfusion image processing more widely accessible. The intended users are perfusion imaging researchers, clinical researchers, and software developers.
(ii) Integrate a testing framework within the repository, so that all code contributions can be easily tested and compared using the same publicly-available test data sets. This will enable researchers and software developers to validate their code, reduce the need to create their own test data and testing frameworks, and enable other researchers to re-use the shared code with greater confidence.
(iii) Leverage the code and testing framework resulting from aims (i) and (ii) to develop a community-led open-source perfusion imaging software package permitting full perfusion processing pipelines to be coded.

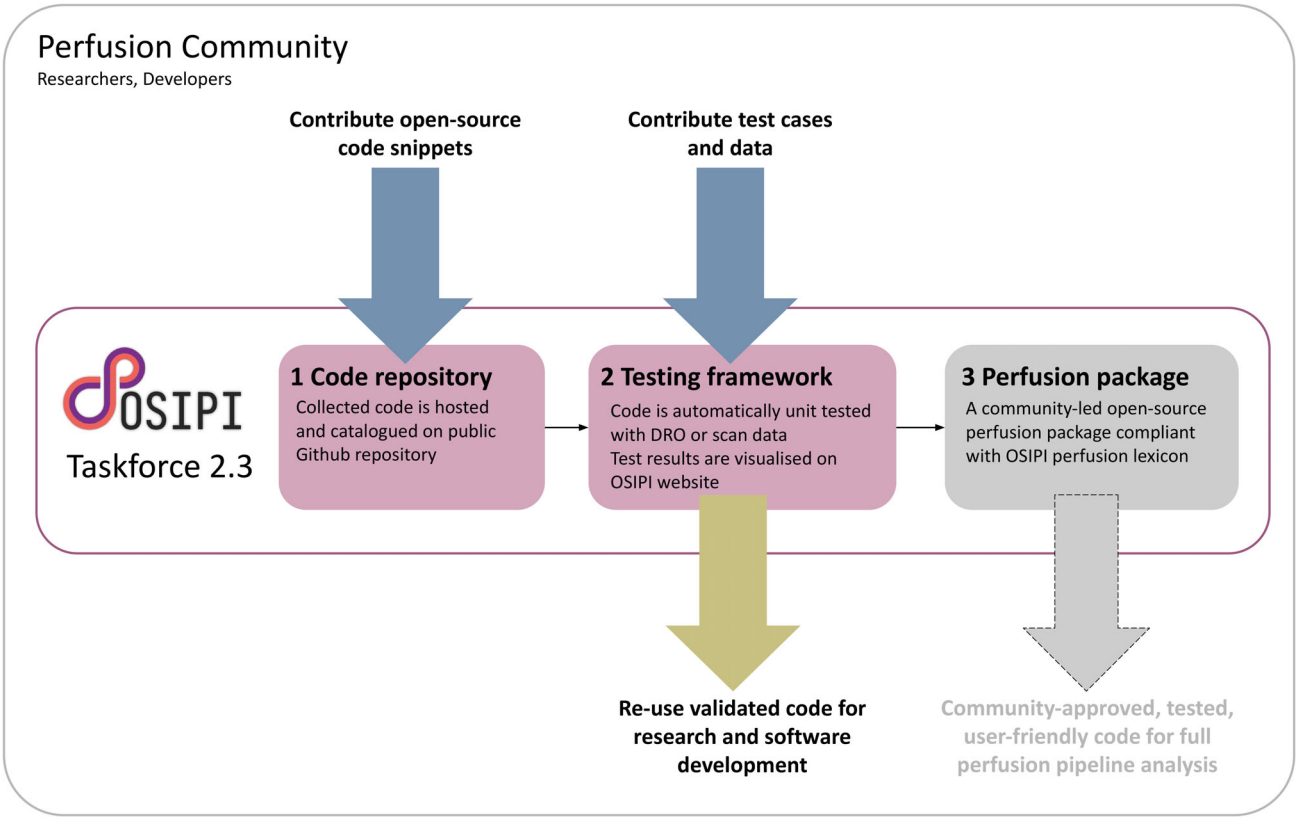

FIGURE 1 Overview of the main aims of OSIPI Taskforce 2.3 and its interactions with the perfusion imaging community. The third aim ("Perfusion package") was not addressed during the initial 2-y cycle.

The purpose of this paper is to report on progress made by the taskforce during the first 2-y cycle of OSIPI. We describe a new framework for collecting, sharing and testing open-source code contributions, summarize the code already shared via the repository, and report (with examples) on the tests implemented across different aspects of functionality. We intend that readers will be encouraged to make use of, contribute to, and join this open-science initiative, and that our approach will inspire initiatives in other fields of medical image processing.

2 | METHODS

2.1 | Taskforce structure and operation

The OSIPI Taskforce 2.3 was established in February 2020 by the OSIPI Executive Management Board. The taskforce consists of a lead (M.J.T.), co-lead (P.v.H.) and members with backgrounds in medical physics, image analysis and research software engineering. Membership and participation are open to all individuals with relevant expertise. Members of the taskforce collaborated via a dedicated *Slack* channel, regular (at least monthly) virtual meetings and the *GitHub* website. The activity of the taskforce followed a roadmap that can be summarized as three phases corresponding to the aims described above: (i) establishment of a code repository (ii) implementation of a testing framework and (iii) development of a community-led perfusion library (Figure 1). Tasks during the first 2-y cycle of OSIPI (2020–2022) were focused on the first two phases, as described below.

2.2 | Scope

The first aim of the taskforce was to lay the foundations for a robust workflow and a sustainable repository that could be extended in the future. Therefore, the initial focus was on code for processing signal-time curves to obtain biophysical parameters. Peripheral steps such as data input/output and obtaining region-of-interest statistics were considered out-of-scope at this stage, since they are not specific to perfusion imaging and accepted software solutions already exist. As open science is one of the key principles of OSIPI, Python was initially targeted as the most popular and well-supported open-source language for scientific computing.

2.3 | Open-source code repository

A *GitHub* repository* was established with the Apache Software License (Version 2.0). Figure 2 gives an overview of the repository structure. Calls for code contributions were made to OSIPI members, the ISMRM Perfusion Study Group, and publicly via OSIPI websites. In addition, individual researchers were approached based on references to Python code in their publications and conference presentations. Contributors were asked to add their code to the repository by creating a feature branch, which was merged with the primary *develop* branch after review. Code contributions were added as a new subfolder within the *src* directory of the repository, labeled according to the originating author and institution. There were no restrictions on the organization, style, or formatting of code within this subfolder. Detailed guidelines on code contribution can be found on the project wiki.†

2.4 | Testing framework

The goals of testing were to identify substantive coding errors and to evaluate and compare the outputs of contributions implementing specific functionalities, using trusted test input data and reference output values. A unit testing framework was implemented using the *pytest* package and *Github Actions*. Test files were created in the *test* directory of the repository and grouped according to function, for example, T_1 mapping and DCE-MRI pharmacokinetic models (Figure 2). Each code contribution was tested using the same input data, reference values and tolerances. The original contributions were not modified for testing, except where essential (for example, adding *__init__.py* files to facilitate import by the test modules). Test data were converted to match the required input format and units before executing the code, and outputs were converted to match the units of the reference values. In some cases, further steps were required in order for the tests to pass. For example, the accuracy of some pharmacokinetic model implementations was increased by interpolating the input time series. For code contributions that implemented a pharmacokinetic model but did not include a fitting routine, the *curve_fit* method from the *scipy* package was used to evaluate the contributed code.[24] In such cases, any additional steps were documented in the test files. A detailed description of the procedure for developing and implementing tests is given in the project wiki.

For each category, at least one set of test data and reference values was included, either simulated (e.g., a digital reference object [DRO]) or based on human in-vivo scans. The aim was to base all tests on publicly available data, software and DROs. Where the dataset itself was not citable, references describing the protocol, patient cohort and the method used to obtain the reference values were cited within the test code. Where necessary, image data were condensed to a limited number of voxels or regions to reduce execution time and computing requirements.

Testing was performed in two stages. First, all tests were automatically executed on a *GitHub* remote runner, triggered by changes within the online repository. For each test case, the code output was compared with the reference output; if the difference between these exceeded the combined absolute and relative tolerances then the test failed and a red "badge" was displayed on the repository home page. The purpose of this testing step was to detect substantial errors in the contributed code or in the test files themselves (e.g., incorrect units for input variables). Therefore, wide tolerance levels were set for these tests in order to detect such errors; these are not intended to indicate an acceptable level of accuracy. Furthermore, we aimed to test the validity of the code and not that of the image acquisition and analysis techniques themselves. Therefore, test cases were avoided for which valid code could not be expected to return accurate results, for example those with low SNR, inadequate temporal resolution and degenerate cases.

Second, a test results website‡ was created to provide end users of the code collection with visual, quantitative representations of the test results. For this purpose, the output values for each of the above test cases and code contributions were exported to comma-separated values files. An automated workflow was established to read and plot these data using *Jupyter* notebooks and the *Jupyter Book* package[25] (Figure 2). The notebooks were exported as HTML pages and can be viewed publicly at a test results website, hosted in a separate repository.§ Results were presented by plotting the deviations of the output values with respect to the reference values (e.g., Bland–Altman plots).

3 | RESULTS

At the time of writing, the taskforce received *Python* source code contributions to the repository (release 1.0.0; https://doi.org/10.5281/zenodo.7729136) comprising 86 implementations of individual perfusion processing steps, contributed by 12 individuals or teams. These include implementations of all core aspects of DCE- and DSC-MRI processing (Table 1). For DCE-MRI, functionality is available for T_1 mapping, bolus arrival time estimation, conversion from signal to concentration, arterial input functions (AIF), pharmacokinetic models and semi-quantitative parameter derivation; for DSC-MRI, functionality is available for conversion from signal to

(A)

Repository: DCE-DSC-MRI_CodeCollection

hosted on: https://github.com/OSIPI/DCE-DSC-MRI_CodeCollection/

```
OSIPI/DCE-DSC-MRI_CodeCollection
├── doc
│    └── code_contributions_record.csv # Overview of code contributions
├── src
│   └── original                       # Original contributed code organized per contribution
│       └── DS_BW_VanderBiltUMC_USA
│       └── JBJA_GUSahlgrenskaSWE
│       └── ...
├── test                               # Python test files, organised by category
│   └── DCEmodels
│       └── data                       # data used for testing stored as csv files
│   └── DSCmodels
│       └── data
│   └── ...
│   └── results-meta.json              # stores meta-information about csv-results files, used by notebooks
├── notebooks                          # Jupyter notebooks and markdown files to build the test-results website
├── LICENCE                            # Apache version 2.0 licence
├── README.md
├── setup.py

Wiki                                   # Information and guidelines
```

(B)

Test-results website: http://osipi.github.io/DCE-DSC-MRI_TestResults

hosted on: https://github.com/OSIPI/DCE-DSC-MRI_TestResults/

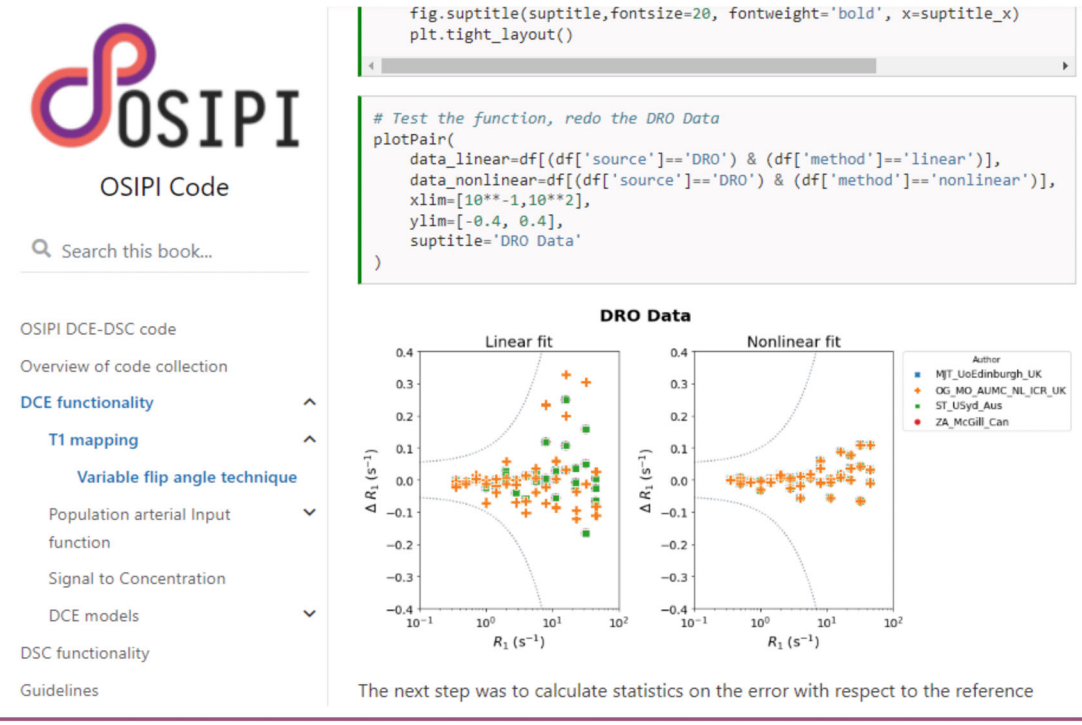

FIGURE 2 Overview of the OSIPI DCE-DSC code repository. (A) The repository directory structure with a description of the content of each directory. The *notebooks* directory contains all files required to publish the results on the *test-results* website. The test results (*csv* format) are automatically pushed to a second repository (DCE-DSC-MRI_TestResults) linked to a website displaying the results. (B) A snapshot of this website.

TABLE 1 Implementations of core perfusion processing functionality collected and tested in release 1.0.0 of the repository

Technique	Processing steps	Implemented methods	Collected	Tested
DCE-MRI	T_1 mapping	Variable flip angle (linear, non-linear, NOVIFAST[26]), DESPOT1-HIFI[27]	11	9
	Bolus arrival time estimation	Piecewise linear quadratic function,[28] estimate delay by fitting Tofts model to first third of the curve	2	0
	Signal to concentration	Conversion from signal to concentration for spoiled gradient echo sequences[a]	8	7
	Arterial input functions	Population functions (Georgiou,[29] Heye,[30] Manning,[31] McGrath,[32] Parker,[33] Wong[34]), patient-specific, parametrization of AIF	15	7
	Pharmacokinetic models	Forward models and linear and non-linear model fitting: Tofts, extended Tofts, Patlak, 2CXM, 2CUM, AATH, steady-state vp, dual-inlet models[35]	41	28
	Parameter derivation	iAUC	1	0
DSC-MRI	Conversion from signal to concentration	Single- and dual- gradient echo	1	0
	Arterial input function	Automated AIF selection	2	0
	Leakage correction	Boxerman–Schmainda–Weiskoff[36]	1	0
	Parameter derivation	Deconvolution; CBV, CBF and MTT estimation	1	1
General		SNR, enhancement ratio estimation	3	0

Note: Some of the collected code contributions were linked to a publication: Bell et al.[37] Berks et al.[38] Johansen O,[39] Mouridsen et al.[40] Orton et al.[41] Rata et al.[42] For more detailed descriptions of the implemented methods, readers are referred to the citations provided in the table.

Abbreviations: 2CUM, two-compartment uptake model; 2CXM, two-compartment exchange model; AATH, adiabatic approximation to the tissue homogeneity; *C*, concentration; CBF, cerebral blood flow; CBV, cerebral blood volume; DESPOT1-HIFI, driven equilibrium single pulse observation of T_1 with high-speed incorporation of RF field inhomogeneities; iAUC, initial area under the curve; MTT, mean transit time; NOVIFAST, non-linear variable flip angle data based T_1 estimator.

[a] Reverse implementations for concentration to signal are also available (not counted separately).

concentration, automatic AIF selection, leakage correction and perfusion parameter derivation. For most categories, multiple contributions implementing the same functionality are available. In many cases, the implementations were mathematically distinct. For example, both linear and non-linear implementations of variable flip angle T_1 mapping are available. Pharmacokinetic models were also implemented using different approaches, including different convolution methods and the use of linear and non-linear least squares fitting routines for parameter estimation. Code contributions also differed according to options and features available: for example, some pharmacokinetic model implementations accepted an artery-capillary delay parameter, while others assumed no delay. Up-to-date descriptions of all collected code are provided as a table in the repository.**

Implementation of the tests for each category of functionality is an ongoing and open-ended process (Table 1). At the time of writing, tests have been implemented for linear and non-linear implementation of variable flip angle T_1 mapping, conversion from signal to concentration for DCE-MRI, population AIFs (Parker, Georgiou, and McGrath), pharmacokinetic models (Patlak, Tofts, extended Tofts-Kety, two-compartment uptake and two-compartment exchange) and DSC-MRI parameter derivation (*CBF* and *CBV*). Table 2 gives an overview of the tests developed for each category of functionality and the tolerances that were used. On the test results website, detailed results of the tests were visualized including relevant background information, and a description of the test data and the tolerances. Example test results are shown in Figure 3, where graphs display the difference between the output and reference values.

For some of the tested functionality categories multiple implementations were available. For T_1 mapping, the four non-linear implementations of variable flip angle T_1 estimation yielded near-identical outputs when processing voxels from the *QIBA T1 DRO*[43] (Figure 3B) and from two sets of in-vivo data. Similar results were obtained for three

TABLE 2 Overview of test data organized by functionality

Technique	Processing steps	Test data description (n = no. of data points)	Reference values [min–max]
DCE-MRI	T_1 mapping	QIBA T1 DRO v3[43]: simulated data ($n = 45$)	R_1: 0.35–45.2 s^{-1}
		In-vivo brain data[44]: voxel data from one patient with mild-stroke based on ROIs drawn in the white matter, deep gray matter and cerebrospinal fluid ($n = 76$)	R_1: 0.14–0.99 s^{-1}
		In-vivo prostate data[45]: cases correspond to randomly selected voxels in the prostate (10 voxels from each of 5 patients with prostate cancer; $n = 50$)	R_1: 0.40–2.78 s^{-1}
	Signal to concentration	In-vivo uterus data[46,47]: cases correspond to randomly selected signal intensity curves from voxels in uterus or aorta from one volunteer ($n = 5$)	$C(t)$: −0.2–5.0 mM
	Parker AIF	The concentration time curve was computed from the AIF parameters given in Table 1 of the original publication[33] using a range of time resolutions, acquisition times and bolus arrival times ($n = 20$)	n.a.
	Georgiou AIF	The concentration values from supplementary material of the original publication were used.[29] A range of temporal resolutions was tested by interpolating the original time series ($n = 7$)	n.a.
	McGrath AIF	The concentration time curve was computed from the AIF parameters given in Table 1 (model B) and equation 5 of the original publication[32] using a range of time resolution, acquisition times and bolus arrival times ($n = 18$)	n.a.
	Tofts model	QIBA DRO (version 11).[43] Averages were taken over large homogeneous patches to get a high-SNR dataset. Noise was added to the high-SNR data to obtain data with different SNRs ($n = 25$)	K^{trans}: 0.05–0.35 min^{-1} v_e: 0.1–0.5
	Extended Tofts model	Anthropomorphic digital MR phantom.[48] Noise was added to the data to obtain data with different SNRs ($n = 15$)	K^{trans}: 0.05–0.08 min^{-1} v_e: 0.15–0.20 v_p: 0.005–0.02
	Patlak model	Concentration-time data generated using Matlab code ($n = 9$)[31]	PS: 0.0–0.15 min^{-1} v_p: 0.1–0.5 $delay$: 0–5 s
	2-compartment exchange model	Concentration-time data generated using Matlab code ($n = 24$)[31]	PS: 0.05–0.15 min^{-1} F_p: 5–40 100 mL/mL/min v_e: 0.1–0.2 v_p: 0.02–0.1 $delay$: 0–5 s
	2-compartment uptake model	Concentration-time data generated using Matlab code. ($n = 51$)[31]	PS: 10^{-5}–0.25 min^{-1} F_p: 5–40 100 mL/mL/min v_p: 0.02–0.1 $delay$: 0–5 s
DSC-MRI	Parameter derivation	Brain DRO consisting of signal time curves representing perfusion scenarios typical of normal gray and white matter ($n = 15$)[49]	CBV = 2–4 mL/100 mL CBF = 5–70 mL/100 mL/min

Note: Publications relating to the test data are cited in the description column.

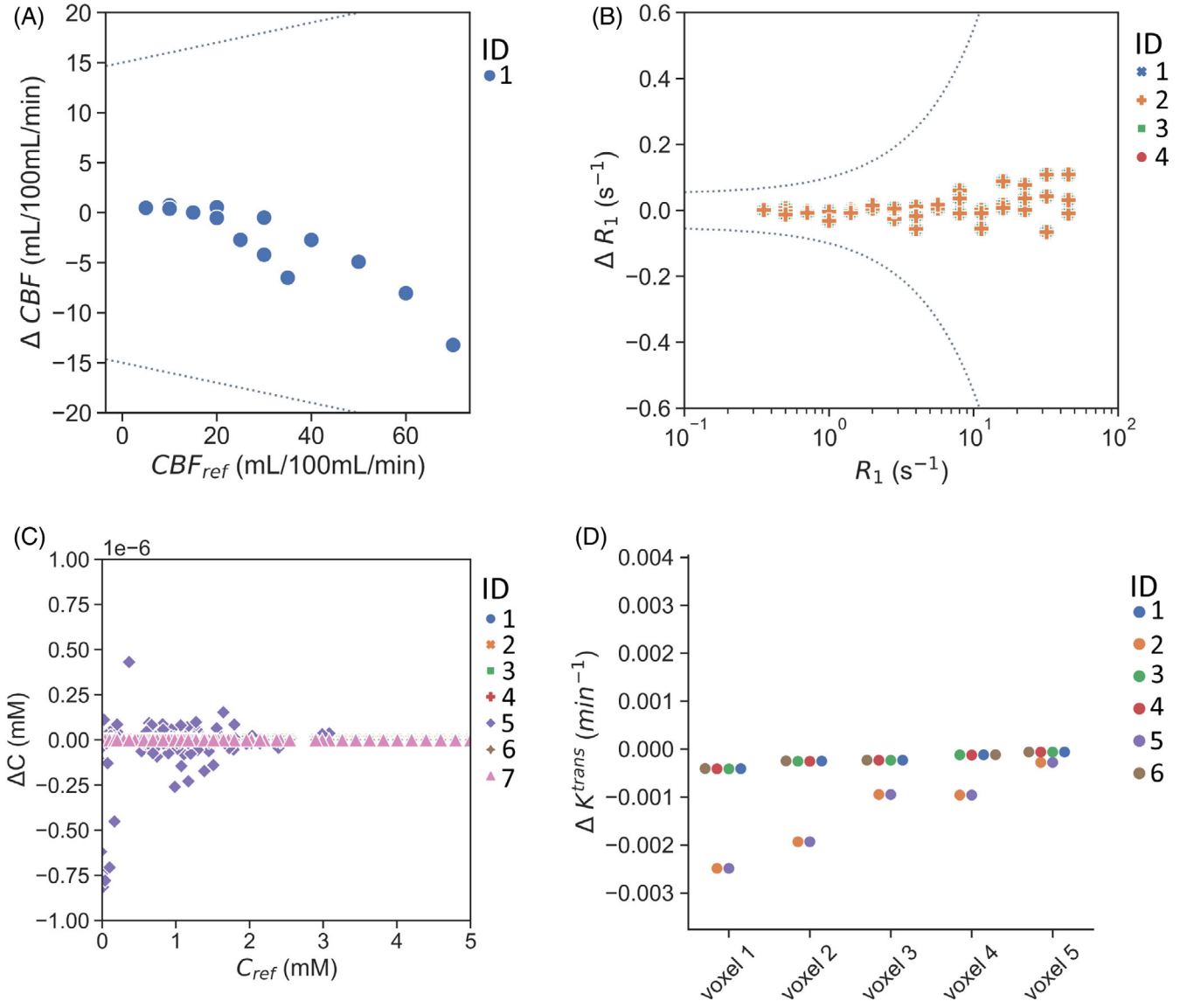

FIGURE 3 Example results from the testing framework. These are snapshots of figures presented on the test-results website: (A) Bland–Altman plot for *CBF* estimation showing the difference between output and reference *CBF* values. At the time of writing one implementation was available. The gray-dashed lines indicate the tolerances used for testing. (B) Bland–Altman plots for variable flip angle T_1 mapping tests using the *QIBA T1 DRO* data. These show the difference between output and reference R_1 values for four different non-linear implementations of T_1 estimation. The gray-dashed lines indicate the tolerances used for testing. Each color represents a different implementation, indicated by a number in the legend. (C) Bland–Altman plot for the conversion of signal intensity to concentration, showing the difference between output and reference concentration values. The tolerances are not shown as they were outside the scale of the plot. (D) Categorical plot for the estimation of K^{Trans} with the Tofts model. These show the difference between output K^{Trans} values and the corresponding reference values for the three test cases with high SNR. Each test case corresponded to a different combination of K^{Trans} and v_e.

linear implementations of variable flip angle T_1 mapping. Seven implementations for the conversion from signal intensity to concentration also resulted in identical output values (Figure 3C), although one implementation showed small deviations (<5e^{-7} mM). For code implementing the Parker population AIF, all implementations resulted in the same concentration-time curve as the published AIF function. However, for code contributions applying an optional time delay to the AIF, some differences were observed when the delay was not a multiple of the temporal resolution. For all five pharmacokinetic models, differences in the fitted parameters were observed between the implementations. For example, six implementations of the Tofts model were tested using data from the *QIBA DRO*,[43] revealing variation in the estimated K^{Trans} and v_e values (Figure 3D).

4 | DISCUSSION AND CONCLUSIONS

In this paper, we described the aims, processes and current status of the OSIPI open-source code repository for DCE- and DSC-MRI processing. The repository constitutes both a resource for the perfusion community to use, and a platform for testing and developing new and existing code.

4.1 | Open-source code collection

The code collection currently includes implementations of the most common steps in DCE- and DSC-MRI analysis pipelines. Most of this code was not publicly available before being contributed. For most categories, multiple implementations of the same functionality are present, which provides opportunities to investigate the impact of differences in software on the reproducibility of quantitative perfusion parameters. In the future, we aim to extend the collection to include additional categories of functionality, and alternative implementations of existing functionality, for example support for other pulse sequences, and parameter estimation based on joint fitting,[50] Bayesian statistics[51] or deep learning.[52–55] Potential DSC-MRI-relevant extensions include multi-echo acquisitions,[56,57] simultaneous spin- and gradient-echo acquisitions to perform vessel size imaging,[58,59] correction for AIF dispersion[60] and partial volume effects,[61] and model based parameter estimation.[62] For DCE-MRI, possible extensions include analyses incorporating finite water exchange rates,[63] other T_1 measurement approaches (for example, saturation-recovery spoiled gradient echo[64]) and patient-specific AIF measurement based on phase or complex signal.[45,65]

A limiting factor was the availability of code in the *Python* language. Other programming languages, such as Matlab, Julia and C++, are also used in the perfusion community. In future, we may include code written in other software languages by translating contributions to *Python* or by using function wrappers. Future calls for code may target specific areas of perfusion functionality and the taskforce may approach specific individuals and groups following literature searches.

4.2 | Testing framework

Unit tests were designed to compare the outputs of different implementations and to provide quality assurance with respect to scientific performance. An automated framework was used, to ensure that testing reflects future updates to the code and any dependencies. This will allow researchers to re-use code written by others with greater confidence. Furthermore, it provides developers with a framework that can be used to validate new software and to compare its scientific performance with that of other implementations. While this work is primarily focused on open-source software, our testing framework can also be run locally by developers of closed-source or commercial software.

There are some limitations of the testing framework. First, ground-truth reference values are often difficult to define. For DROs, the ability of software to match the reference values depends on the algorithms used to generate the data, the simulated imaging protocol (for example, the temporal resolution) and other factors. For in-vivo data, the reference values depend on the code used to process the data and will be influenced by noise, artifacts, and the processing strategy. Thus, the testing framework, while designed to assure an acceptable level of quality and to compare the quantitative outputs of different code contributions, is not intended as a means to rank or recommend specific contributions. Indeed, there may not be a single implementation suitable for all applications and use cases. For example, in the case of pharmacokinetic model implementations, there were substantial differences in the contributed algorithms: some estimated tracer concentration via a simple discrete convolution of the AIF and impulse response function, while others used more exact approximations to the convolution integral; in one set of implementations, the AIF was parameterized so that concentrations could be calculated analytically. It is expected that the nature of the underlying algorithm will affect both the accuracy and the computational efficiency, dependent on the use case: for example, the simple convolution approach may be accurate and fast only if the impulse response function and the AIF are sampled with sufficient temporal resolution. Therefore, it is left to the user to select an implementation appropriate to their needs and to review the underlying methodology. The current testing framework should nevertheless aid future initiatives to standardize and harmonize perfusion processing.

Secondly, adding new code contributions to the current testing framework requires some manual intervention by taskforce members or the user themselves. In the future it would be beneficial to automate this process. Thirdly, speed and robustness (for example, to invalid inputs) were not assessed at this stage. Nor did we review the code in a line-by-line manner to detect errors (although we note that our testing framework did detect a small number of coding errors that were corrected with permission from the contributing authors). Fourthly, the current test data does not cover parameter ranges relevant to all applications. For example, our extended Tofts model data has a maximum K^{trans} value of 0.08 min^{-1}, which is below the values seen

in many tumors. However, as a community-led initiative, OSIPI welcomes contributions of additional test data from members of the perfusion community. Finally, we focused on test cases where there is a reasonable expectation of obtaining a valid result. For example, when testing pharmacokinetic model implementations, we used test data that was generated and tested using the same model, that had a high temporal resolution and that yielded a single well-defined set of model parameters; it was not our aim to replicate the extensive literature on the validity and interpretations of such models as a function of tissue biology and experimental parameters.

4.3 | Outlook

The OSIPI DCE-DSC code repository is an ongoing project that welcomes new members, code contributions and test data from the perfusion community. In future, we will extend the range of functionality within the repository and extend the testing framework to cover additional functionality and use cases.

A longer-term objective of the taskforce is to harmonize and integrate code contributions into a coherent code library that is validated, user-friendly and based on community-consensus methodology. The library shall also dovetail with the OSIPI contrast-agent based perfusion MRI lexicon (Taskforce 4.2), such that quantities, units, models and processes referenced in the library correspond precisely to consensus definitions. This will support standardized processing, transparent reporting and ease of replication. The development of such a library depends on the availability of funding, research software engineering expertise and perfusion community engagement. However, the repository and the code therein provide a foundation for future collaborative software development.

In conclusion, we have presented a community-led model for code sharing and testing. By facilitating the re-use of tested code and the benchmarking of new code, we expect that the OSIPI DCE-DSC code repository will be a valuable resource to researchers and developers. The repository should be of particular benefit to new researchers in the field who will not need to begin coding from scratch. We hope that this will result in improved reproducibility, reduced duplicate development, and support the wider use of perfusion imaging as endpoints in clinical trials.

AFFILIATIONS

[1]Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

[2]Hyperfine Inc, Clinical Science, Guilford, USA

[3]Quantitative Biomedical Imaging Laboratory, Division of Cancer Sciences, The University of Manchester, Manchester, UK

[4]Corewell Health William Beaumont University Hospital, Diagnostic Radiology, Royal Oak, USA

[5]Edinburgh Imaging and Centre for Cardiovascular Science, University of Edinburgh, Edinburgh, UK

[6]Department of Radiology and Nuclear Medicine, Amsterdam UMC location University of Amsterdam, Amsterdam, The Netherlands

[7]Cancer Center Amsterdam, Imaging and Biomarkers, Amsterdam, The Netherlands

[8]Institute of Medical Physics, The University of Sydney, Sydney, Australia

[9]Department of Medical Radiation Sciences, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[10]Department of Medical Physics and Biomedical Engineering, Sahlgrenska University Hospital, Gothenburg, Sweden

[11]Aarhus University Hospital, Danish Centre for Particle Therapy, Aarhus, Denmark

[12]Aarhus University, Department of Clinical Medicine, Aarhus, Denmark

[13]Division of Informatics, Imaging, and Data Science, School of Health Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK

[14]Geoffrey Jefferson Brain Research Centre, Manchester Academic Health Science Centre, Northern Care Alliance NHS Group, The University of Manchester, Manchester, UK

[15]MR Research Collaborations, Siemens Healthcare Pty Ltd, Melbourne, Australia

[16]Genentech, Inc, Clinical Imaging Group, South San Francisco, USA

[17]University of Sheffield, Department of Infection, Immunity and Cardiovascular Disease, Sheffield, UK

[18]University of Edinburgh, Edinburgh Imaging and Centre for Clinical Brain Sciences, Edinburgh, UK

ACKNOWLEDGMENTS

We acknowledge the OSIPI membership and leadership for their suggestions and feedback, Ole Gunnar Johansen (University of Oslo), Matthew Orton (The Royal Marsden NHS Foundation Trust), Brian Welch (Philips Healthcare) and David Smith (Institute of Imaging Science, Vanderbilt University) for contributing code, Joanna M. Wardlaw (University of Edinburgh) for providing test data, and Uulke van der Heide (the Netherlands Cancer Institute) for valuable feedback on the manuscript.

CONFLICT OF INTEREST STATEMENT

Dr Sudarshan Ragunathan is employed by Hyperfine Inc. Dr Simon Lévy is employed by Siemens Healthcare Pty Ltd. Dr Laura Bell is employed by Genentech, Inc.

DATA AVAILABILITY STATEMENT

Latest versions of the code, test data, and test results are openly available at https://github.com/OSIPI/DCE-DSC-MRI_CodeCollection; a website displaying the test

results graphically is hosted in a second repository: https://github.com/OSIPI/DCE-DSC-MRI_TestResults. The version (1.0.0) of the repository described in this article is available at https://doi.org/10.5281/zenodo.7729136.

ENDNOTES

*https://github.com/OSIPI/DCE-DSC-MRI_CodeCollection
†https://github.com/OSIPI/DCE-DSC-MRI_CodeCollection/wiki/
‡https://osipi.github.io/DCE-DSC-MRI_TestResults/intro.html
§https://github.com/OSIPI/DCE-DSC-MRI_TestResults
**https://osipi.github.io/DCE-DSC-MRI_TestResults/overview_of_code_collection.html

ORCID

Petra J van Houdt https://orcid.org/0000-0001-7431-8386
Sudarshan Ragunathan https://orcid.org/0000-0002-2871-1758
Michael Berks https://orcid.org/0000-0003-4727-2006
Lucy E Kershaw https://orcid.org/0000-0001-9510-7940
Oliver J Gurney-Champion https://orcid.org/0000-0003-1750-6617
Sirisha Tadimalla https://orcid.org/0000-0003-2054-055X
Jonathan Arvidsson https://orcid.org/0000-0002-6955-6360
Yu Sun https://orcid.org/0000-0002-5206-8782
Jesper Kallehauge https://orcid.org/0000-0003-3705-5390
Michael J Thrippleton https://orcid.org/0000-0001-7858-9917

REFERENCES

1. Cao Y. The promise of dynamic contrast-enhanced imaging in radiation therapy. *Semin Radiat Oncol*. 2011;21:147-156. doi:10.1016/j.semradonc.2010.11.001
2. Sung YS, Park B, Choi Y, et al. Dynamic contrast-enhanced MRI for oncology drug development. *J Magn Reson Imaging*. 2016;44:251-264. doi:10.1002/jmri.25173
3. Raja R, Rosenberg GA, Caprihan A. MRI measurements of blood–brain barrier function in dementia: a review of recent studies. *Neuropharmacology*. 2018;134:259-271. doi:10.1016/j.neuropharm.2017.10.034
4. Thrippleton MJ, Backes WH, Sourbron S, et al. Quantifying blood–brain barrier leakage in small vessel disease: review and consensus recommendations. *Alzheimers Dement*. 2019;15:840-858. doi:10.1016/j.jalz.2019.01.013
5. Duering M, Biessels GJ, Brodtmann A, et al. Neuroimaging standards for research into small vessel disease—advances since 2013. *Lancet Neurol*. 2023;22:602-618. doi:10.1016/S1474-4422(23)00131-X
6. O'Connor JPB, Aboagy EO, Adams JE, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol*. 2016;14:169-186. doi:10.1038/nrclinonc.2016.162
7. Kim H. Variability in quantitative DCE-MRI: sources and solutions. *J Nat Sci*. 2018;4:e484.
8. Shukla-Dave A, Obuchowski NA, Chenevert TL, et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *J Magn Reson Imaging*. 2019;49:e101-e121. doi:10.1002/jmri.26518
9. Boxerman JL, Quarles CC, Hu LS, et al. Consensus recommendations for a dynamic susceptibility contrast MRI protocol for use in high-grade gliomas. *Neuro Oncol*. 2020;22:1262-1275. doi:10.1093/neuonc/noaa141
10. Ellingson BM, Bendszus M, Boxerman J, et al. Consensus recommendations for a standardized brain tumor imaging protocol in clinical trials. *Neuro Oncol*. 2015;17:1188-1198. doi:10.1093/neuonc/nov095
11. Ger RB, Mohamed ASR, Awan MJ, et al. A multi-institutional comparison of dynamic contrast-enhanced magnetic resonance imaging parameter calculations. *Sci Rep*. 2017;7:11185. doi:10.1038/s41598-017-11554-w
12. Delacoste E, Delattre B, Wanyanga P, Vargas MI. Comparing dynamic susceptibility contrast perfusion post-processing with different clinically available software among patients affected of a high-grade glioma. *J Neuroradiol*. 2022;49:412-420. doi:10.1016/j.neurad.2020.09.010
13. Heye T, Davenport MS, Horvath JJ, et al. Reproducibility of dynamic contrast-enhanced MR imaging. Part I. perfusion characteristics in the female pelvis by using multiple computer-aided diagnosis perfusion analysis solutions. *Radiology*. 2013;266:801-811. doi:10.1148/radiol.12120278
14. Kudo K, Christensen S, Sasaki M, et al. Accuracy and reliability assessment of CT and MR perfusion analysis software using a digital phantom. *Radiology*. 2013;267:201-211. doi:10.1148/radiol.12112618
15. Huang W, Li X, Chen Y, et al. Variations of dynamic contrast-enhanced magnetic resonance imaging in evaluation of breast cancer therapy response: a multicenter data analysis challenge. *Transl Oncol*. 2014;7:153-166. doi:10.1593/tlo.13838
16. Beuzit L, Eliat PA, Brun V, et al. Dynamic contrast-enhanced MRI: study of inter-software accuracy and reproducibility using simulated and clinical data. *J Magn Reson Imaging*. 2016;43:1288-1300. doi:10.1002/jmri.25101
17. Bell LC, Semmineh N, An H, et al. Evaluating multisite rCBV consistency from DSC-MRI imaging protocols and Postprocessing software across the NCI quantitative imaging network sites using a digital reference object (DRO). *Tomography*. 2019;5:110-117. doi:10.18383/j.tom.2018.00041
18. Prabhu P, Zhang Y, Ghosh S, et al. A survey of the practice of computational science. *State of the Practice Reports on – SC'11*. ACM Press; 2011:1. doi:10.1145/2063348.2063374
19. Hannay JE, MacLeod C, Singer J, Langtangen HP, Pfahl D, Wilson G. How do scientists develop and use scientific software? *2009 ICSE Workshop on Software Engineering for Computational Science and Engineering*. IEEE; 2009:1-8. doi:10.1109/SECSE.2009.5069155
20. Wilson G, Aruliah DA, Brown CT, et al. Best practices for scientific computing. *PLoS Biol*. 2014;12:e1001745. doi:10.1371/journal.pbio.1001745

21. Wilson G, Bryan J, Cranston K, Kitzes J, Nederbragt L, Teal TK. Good enough practices in scientific computing. *PLoS Comput Biol*. 2017;13:e1005510. doi:10.1371/journal.pcbi.1005510
22. Morton L. Uphold the code: How complete, detailed and Open code can enhance understanding, improve reproducibility, and change the shape of the research article. https://theplosblog.plos.org/2022/05/uphold-the-code
23. Bell LC, Suzuki Y, van Houdt PJ, Sourbron S, Mutsaerts HJMM. The road to the ISMRM OSIPI: a community-led initiative for reproducible perfusion MRI. *Magn Reson Med*. 2023. doi:10.1002/mrm.29736
24. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in python. *Nat Methods*. 2020;17:261-272. doi:10.1038/s41592-019-0686-2
25. Executable Books Community. Jupyter Book. 2020. doi:10.5281/zenodo.4539666
26. Ramos-Llorden G, Vegas-Sanchez-Ferrero G, Bjork M, et al. NOVIFAST: a fast algorithm for accurate and precise VFA MRI T_1 mapping. *IEEE Trans Med Imaging*. 2018;37:2414-2427. doi:10.1109/TMI.2018.2833288
27. Deoni SCL. High-resolution T_1 mapping of the brain at 3 T with driven equilibrium single pulse observation of T_1 with high-speed incorporation of RF field inhomogeneities (DESPOT1-HIFI). *J Magn Reson Imaging*. 2007;26:1106-1111. doi:10.1002/jmri.21130
28. Cheong LH, Koh TS, Hou Z. An automatic approach for estimating bolus arrival time in dynamic contrast MRI using piecewise continuous regression models. *Phys Med Biol*. 2003;48:N83-N88. doi:10.1088/0031-9155/48/5/403
29. Georgiou L, Wilson DJ, Sharma N, Perren TJ, Buckley DL. A functional form for a representative individual arterial input function measured from a population using high temporal resolution DCE MRI. *Magn Reson Med*. 2019;81:1955-1963. doi:10.1002/mrm.27524
30. Heye AK, Thrippleton MJ, Armitage PA, et al. Tracer kinetic modelling for DCE-MRI quantification of subtle blood–brain barrier permeability. *Neuroimage*. 2016;125:446-455. doi:10.1016/j.neuroimage.2015.10.018
31. Manning C, Stringer M, Dickie B, et al. Sources of systematic error in DCE-MRI estimation of low-level blood–brain barrier leakage. *Magn Reson Med*. 2021;86:1888-1903. doi:10.1002/mrm.28833
32. McGrath DM, Bradley DP, Tessier JL, Lacey T, Taylor CJ, Parker GJM. Comparison of model-based arterial input functions for dynamic contrast-enhanced MRI in tumor bearing rats. *Magn Reson Med*. 2009;61:1173-1184. doi:10.1002/mrm.21959
33. Parker GJM, Roberts C, Macdonald A, et al. Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI. *Magn Reson Med*. 2006;56:993-1000. doi:10.1002/mrm.21066
34. Wong KH, Panek R, Welsh L, et al. The predictive value of early assessment after 1 cycle of induction chemotherapy with 18F-FDG PET/CT and diffusion-weighted MRI for response to radical Chemoradiotherapy in head and neck squamous cell carcinoma. *J Nucl Med*. 2016;57:1843-1850. doi:10.2967/jnumed.116.174433
35. Sourbron SP, Buckley DL. Classic models for dynamic contrast-enhanced MRI. *NMR Biomed*. 2013;26:1004-1027. doi:10.1002/nbm.2940
36. Boxerman JL, Schmainda KM, Weisskoff RM. Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not. *AJNR Am J Neuroradiol*. 2006;27:859-867. http://www.ncbi.nlm.nih.gov/pubmed/16611779
37. Bell LC, Ragunathan S, Fathi Kazerooni A. Contrast agent-based perfusion MRI methods. In: Choi IY, Jezzard P, eds. *Advanced Neuro MR Techniques and Applications*. Academic Press; 2021:195-209. doi:10.1016/B978-0-12-822479-3.00024-5
38. Berks M, Parker G, Little R, Cheung S. Madym: a C++ toolkit for quantitative DCE-MRI analysis. *J Open Source Softw*. 2021;6:3523. doi:10.21105/joss.03523
39. Johansen OG. *DCE-MRI Pharmacokinetic Model Optimization and Implications for Brain Cancer Imaging*. University of Oslo; 2018.
40. Mouridsen K, Christensen S, Gyldensted L, Østergaard L. Automatic selection of arterial input function using cluster analysis. *Magn Reson Med*. 2006;55:524-531. doi:10.1002/mrm.20759
41. Orton MR, d'Arcy JA, Walker-Samuel S, et al. Computationally efficient vascular input function models for quantitative kinetic modelling using DCE-MRI. *Phys Med Biol*. 2008;53:1225-1239. doi:10.1088/0031-9155/53/5/005
42. Rata M, Collins DJ, Darcy J, et al. Assessment of repeatability and treatment response in early phase clinical trials using DCE-MRI: comparison of parametric analysis using MR- and CT-derived arterial input functions. *Eur Radiol*. 2016;26:1991-1998. doi:10.1007/s00330-015-4012-9
43. Barboriak D. DCE-MRI DRO data. https://qidw.rsna.org/#collection/594810551cac0a4ec8ffe574/folder/5e20cffeb3467a6a9212f87f
44. Clancy U, Garcia DJ, Stringer MS, et al. Rationale and design of a longitudinal study of cerebral small vessel diseases, clinical and imaging outcomes in patients presenting with mild ischaemic stroke: mild stroke study 3. *Eur Stroke J*. 2021;6:81-88. doi:10.1177/2396987320929617
45. Klawer EME, van Houdt PJ, Simonis FFJ, et al. Improved repeatability of dynamic contrast-enhanced MRI using the complex MRI signal to derive arterial input functions: a test–retest study in prostate cancer patients. *Magn Reson Med*. 2019;81:3358-3369. doi:10.1002/mrm.27646
46. UCON. https://www.birmingham.ac.uk/research/bctu/trials/womens/ucon/ucon-home.aspx.
47. Reavey JJ, Walker C, Nicol M, et al. Markers of human endometrial hypoxia can be detected in vivo and ex vivo during physiological menstruation. *Hum Reprod*. 2021;36:941-950. doi:10.1093/humrep/deaa379
48. Bosca RJ, Jackson EF. Creating an anthropomorphic digital MR phantom—An extensible tool for comparing and evaluating quantitative imaging algorithms. *Phys Med Biol*. 2016;61:974-982. doi:10.1088/0031-9155/61/2/974
49. Chakwizira A, Ahlgren A, Knutsson L, Wirestam R. Non-parametric deconvolution using Bézier curves for quantification of cerebral perfusion in dynamic susceptibility contrast MRI. *MAGMA*. 2022;35:791-804. doi:10.1007/s10334-021-00995-0

50. Dickie BR, Banerji A, Kershaw LE, et al. Improved accuracy and precision of tracer kinetic parameters by joint fitting to variable flip angle and dynamic contrast enhanced MRI data. *Magn Reson Med*. 2016;76:1270-1281. doi:10.1002/mrm.26013
51. Mittermeier A, Ertl-Wagner B, Ricke J, Dietrich O, Ingrisch M. Bayesian pharmacokinetic modeling of dynamic contrast-enhanced magnetic resonance imaging: validation and application. *Phys Med Biol*. 2019;64:18NT02. doi:10.1088/1361-6560/ab3a5a
52. Ulas C, Das D, Thrippleton MJ, et al. Convolutional neural networks for direct inference of pharmacokinetic parameters: application to stroke dynamic contrast-enhanced MRI. *Front Neurol*. 2019;9. doi:10.3389/fneur.2018.01147
53. Zou J, Balter JM, Cao Y. Estimation of pharmacokinetic parameters from DCE-MRI by extracting long and short time-dependent features using an LSTM network. *Med Phys*. 2020;47:3447-3457. doi:10.1002/mp.14222
54. Fang K, Wang Z, Li Z, et al. Convolutional neural network for accelerating the computation of the extended Tofts model in dynamic contrast-enhanced magnetic resonance imaging. *J Magn Reson Imaging*. 2021;53:1898-1910. doi:10.1002/jmri.27495
55. Ottens T, Barbieri S, Orton MR, et al. Deep learning DCE-MRI parameter estimation: application in pancreatic cancer. *Med Image Anal*. 2022;80:102512. doi:10.1016/j.media.2022.102512
56. Paulson ES, Prah DE, Schmainda KM. Spiral perfusion imaging with consecutive echoes (SPICE™) for the simultaneous mapping of DSC- and DCE-MRI parameters in brain tumor patients: theory and initial feasibility. *Tomography*. 2016;2:295-307. doi:10.18383/j.tom.2016.00217
57. Quarles CC, Gore JC, Xu L, Yankeelov TE. Comparison of dual-echo DSC-MRI- and DCE-MRI-derived contrast agent kinetic parameters. *Magn Reson Imaging*. 2012;30:944-953. doi:10.1016/j.mri.2012.03.008
58. Kiselev VG, Strecker R, Ziyeh S, Speck O, Hennig J. Vessel size imaging in humans. *Magn Reson Med*. 2005;53:553-563. doi:10.1002/mrm.20383
59. Schmiedeskamp H, Straka M, Newbould RD, et al. Combined spin- and gradient-echo perfusion-weighted imaging. *Magn Reson Med*. 2012;68:30-40. doi:10.1002/mrm.23195
60. Mehndiratta A, Calamante F, MacIntosh BJ, Crane DE, Payne SJ, Chappell MA. Modeling and correction of bolus dispersion effects in dynamic susceptibility contrast MRI. *Magn Reson Med*. 2014;72:1762-1774. doi:10.1002/mrm.25077
61. Ahlgren A, Wirestam R, Lind E, Ståhlberg F, Knutsson L. A linear mixed perfusion model for tissue partial volume correction of perfusion estimates in dynamic susceptibility contrast MRI: impact on absolute quantification, repeatability, and agreement with pseudo-continuous arterial spin labeling. *Magn Reson Med*. 2017;77:2203-2214. doi:10.1002/mrm.26305
62. Mouridsen K, Friston K, Hjort N, Gyldensted L, Østergaard L, Kiebel S. Bayesian estimation of cerebral perfusion using a physiological model of microvasculature. *Neuroimage*. 2006;33:570-579. doi:10.1016/j.neuroimage.2006.06.015
63. Dickie BR, Parker GJM, Parkes LM. Measuring water exchange across the blood–brain barrier using MRI. *Prog Nucl Magn Reson Spectrosc*. 2020;116:19-39. doi:10.1016/j.pnmrs.2019.09.002
64. Wennen M, Moll M, Marcus T, et al. A model for fat-suppressed variable flip angle T_1-mapping and dynamic contrast enhanced MRI. *Proceedings of the ISMRM-ESMRMB & ISMRT 31th Annual Meeting*. International Society for Magnetic Resonance in Medicine; 2022:2693.
65. Garpebring A, Wirestam R, Yu J, Asklund T, Karlsson M. Phase-based arterial input functions in humans applied to dynamic contrast-enhanced MRI: potential usefulness and limitations. *MAGMA*. 2011;24:233-245. doi:10.1007/s10334-011-0257-8

How to cite this article: van Houdt PJ, Ragunathan S, Berks M, et al. Contrast-agent-based perfusion MRI code repository and testing framework: ISMRM Open Science Initiative for Perfusion Imaging (OSIPI). *Magn Reson Med*. 2023;1-13. doi: 10.1002/mrm.29826